

## PATENT ABSTRACTS OF JAPAN

(11)Publication number : 2002-152755

(43)Date of publication of application : 24.05.2002

(51)Int.Cl.

H04N 7/32

H03M 7/30

(21)Application number : 2000-341354

(71)Applicant : NIPPON TELEGR &amp; TELEPH CORP &lt;NTT&gt;

(22)Date of filing : 09.11.2000

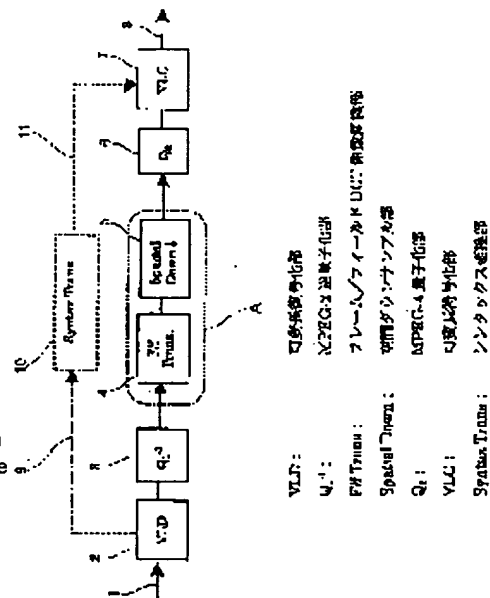
(72)Inventor : TSUJI HIROYUKI  
YASHIMA YOSHIYUKI  
SAKATANI TORU  
KOBAYASHI NAOKI

(54) TRANSCODE METHOD BETWEEN VIDEO STREAMS, AND TRANSCODER, AND RECORDING MEDIUM WITH PROGRAM FOR THE TRANSCODE METHOD RECORDED THEREON

(57)Abstract:

PROBLEM TO BE SOLVED: To attain re-coding of a stream of a different kind by reducing a transcode processing time.

SOLUTION: A variable length decoding section 2 decodes an input video stream 1 subjected to compression coding through the combination of motion compensation and transform coding and extracts a syntax parameter from the stream 1. An input inverse quantization section 3 applies inverse quantization to transform coefficients of each block of the video stream, a pre-processing transform section 4 and a spatial down-sample section 5 obtain a transform coefficient to transform the video stream into a frame image with different image or special resolution, an output quantization section 6 again applies quantization to the transform coefficient, and a variable length coding section 7 applies variable length coding to an output of the quantization section 6. Data from a syntax transform section 10 are given to the variable length coding section 7, where the data are composited with the transform coefficient outputted from the output quantization section 6 and the variable length coding section 7 finally generates and outputs and output video stream.



## LEGAL STATUS

[Date of request for examination]

[Date of sending the examiner's decision of rejection]

[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]

[Date of final disposal for application]

[Patent number]

[Date of registration]

[Number of appeal against examiner's decision of rejection]

[Date of requesting appeal against examiner's decision of rejection]

[Date of extinction of right]

Copyright (C); 1998,2003 Japan Patent Office

\* NOTICES \*

Japan Patent Office is not responsible for any damages caused by the use of this translation.

1. This document has been translated by computer. So the translation may not reflect the original precisely.
2. \*\*\*\* shows the word which can not be translated.
3. In the drawings, any words are not translated.

---

CLAIMS

---

[Claim(s)]

[Claim 1] An input image stream by which compression coding was carried out with a hybrid code-ized method which combined a motion compensation and conversion coding. It is the transformer code method changed and outputted at other image streams from which syntax differs. A syntax parameter required for transformer codes including input image stream lost-motion vector information is extracted. An extracted parameter is changed and saved so that syntax of an output stream may be suited. Without performing inverse transformation to a transform coefficient if needed, thin out a bidirectional prediction frame from an input image stream, and time amount resolution is reduced. Reduce image size of this image stream that time amount resolution reduced, without performing inverse transformation to a transform coefficient, and space resolution is reduced. A transformer code method between image streams characterized by what an image stream from which the time of an input, space resolution, or time amount resolution differs is generated and outputted for under right syntax from this acquired transform coefficient information and this saved syntax parameter information.

[Claim 2] A transformer code method between image streams characterized by amending image deterioration of this image stream that time amount and space resolution reduced in a transformer code method according to claim 1, without performing inverse transformation to a transform coefficient.

[Claim 3] A transformer coder which changes and outputs an input image stream by which compression coding was carried out with a hybrid code-ized method which is characterized by providing the following, and which combined a motion compensation and conversion coding at other image streams from which syntax differs. A means to extract a syntax parameter required for transformer codes including input image stream lost-motion vector information. A means to change and save an extracted parameter so that syntax of an output stream may be suited. A means to thin out a bidirectional prediction frame from an input image stream, and to reduce time amount resolution, without performing inverse transformation to a transform coefficient if needed. A means which generates and outputs an image stream from which the time of an input, space resolution, or time amount resolution differs under right syntax from a means to reduce image size of this image stream that time amount resolution reduced, and to reduce space resolution, without performing inverse transformation to a transform coefficient, and this acquired transform coefficient information and this saved syntax parameter information.

[Claim 4] A transformer coder characterized by having a means to amend image deterioration of this image stream that time amount and space resolution reduced in a transformer coder according to claim 1, without performing inverse transformation to a transform coefficient.

[Claim 5] An input image stream by which compression coding was carried out with a hybrid code-ized method which combined a motion compensation and conversion coding. It is the record medium which recorded a program of a transformer code method changed and outputted on other image streams from which syntax differs. A process in which a syntax parameter required for transformer codes including input image stream lost-motion vector information is extracted, A process in which an extracted parameter is changed and saved so that syntax of an output stream may be suited, A process in which

thin out a bidirectional prediction frame from an input image stream, and time amount resolution is reduced, without performing inverse transformation to a transform coefficient if needed, A process in which reduce image size of this image stream that time amount resolution reduced, and space resolution is reduced, without performing inverse transformation to a transform coefficient, A process which generates and outputs an image stream from which the time of an input, space resolution, or time amount resolution differs under right syntax from this acquired transform coefficient information and this saved syntax parameter information, A record medium which recorded a program of a transformer code method characterized by recording a program for performing a computer on a record medium which this computer can read.

[Claim 6] A record medium which recorded a program of a transformer code method characterized by having a process which amends image deterioration of this image stream that time amount and space resolution reduced in a record medium according to claim 5, without performing inverse transformation to a transform coefficient.

---

[Translation done.]

## \* NOTICES \*

Japan Patent Office is not responsible for any damages caused by the use of this translation.

- 1.This document has been translated by computer. So the translation may not reflect the original precisely.
- 2.\*\*\*\* shows the word which can not be translated.
- 3.In the drawings, any words are not translated.

---

DETAILED DESCRIPTION

---

[Detailed Description of the Invention]

[0001]

[The technical field to which invention belongs] This invention belongs to image coding technology, and relates to the method and transformer coder which carry out recoding (transformer code) of the image data already encoded with a certain bit rate to the image data of a lower bit rate with which time amount differs from space resolution.

[0002]

[Description of the Prior Art] There is a precedent of (1) reference "Susie J.We, John G.Apostolopoulos, NickFeamster, "Field-to-Frame Transcoding with Spatial and Temporal Downsampling", ICIP'99.1999" etc. in the method of carrying out recoding of the image data by which compression coding was carried out by the hybrid code-ized method which combined a motion compensation and conversion coding to the image data of a low bit rate with which time amount differs from space resolution.

[0003] On the other hand to the method of performing recoding of image data, without performing inverse transformation to a transform coefficient (2) -- reference "Nishimura \*\*, Hiroyuki Kasai, Kazuyuki Takaya, the Kameyama \*\*, and sakaki -- Naoaki -- Hanamura \*\*, the Tominaga \*\*\*\*, "development of real-time MPEG-2 video transformer coder software", \*\*\*\*\*, vol.AVM 25-5, pp.25-30.1999", (3) Reference "Pedro A.A.Assuncao, Mohammed Ghanbari, and "A Frequency-Domain Video Transcoder for Dynamic Bit-Rate There are precedents, such as Reduction of MPEG-2 Bit Streams", IEEETrans.CSVT, vol.8, pp.953-967, and Dec.1998."

[0004]

[Problem(s) to be Solved by the Invention] By the above-mentioned method of (1), while reducing image size in a pixel field once elongating the compressed image data by inverse transformation in order to perform recoding to the image data with which time amount differs from space resolution, in order to take again the redundant configuration of performing coding by the motion compensation and conversion coding, there was a problem that the transformer code processing time increased.

[0005] Moreover, the proposal (above (2) and (3)) of a conversion method without the inverse transformation aiming at compaction of the processing time accompanying such a transformer code is restricted to the transformer code into which the application field does not change time amount and space resolution between the streams of MPEG-2 ->MPEG-2 grade of the same kind, and it was not proposed about recoding to the different-species stream accompanied by reduction of time amount or space resolution.

[0006] The purpose of this invention is to offer the record medium which recorded the program of the transformer code method on the transformer code method and transformer coder list which shorten the transformer code processing time and can perform recoding to a different-species stream.

[0007]

[Means for Solving the Problem] A means which is expressed below is used for this invention in order to solve the above-mentioned technical problem.

[0008] First, a functional-block group which unifies and reduces a transform coefficient changed for every block, without performing inverse transformation is prepared. Image data inputted into this equipment is inputted into this functional-block group as a transform coefficient block through variable-length decode and reverse quantization. Here, space resolution reduction of the whole image is realized by generating the new block which unified two or more adjoining blocks, without reducing a block size. This functional-block group is \*\*\*\*\* (ed), it is re- (usually coarse) quantized according to an aim bit rate, and image data to which space resolution was reduced creates an image stream finally outputted from this equipment by adding and compounding required syntax data.

[0009] By such processing, inverse transformation of a transform coefficient is not carrying out once, and shortens the total processing time.

[0010] Moreover, since image deterioration accompanying a drift error generally is not avoided about re-quantization, it enables it to add the feedback loop which amends a drift error if needed with frequency data in this invention. A functional-block group which performs a motion compensation with frequency data is contained in this feedback loop, and image quality after a transformer code is raised.

[0011] From the above thing, this invention is characterized by record medium which recorded a transformer code method on the following transformer code methods and transformer coder lists.

[0012] (Invention of a transformer code method) An input image stream by which compression coding was carried out with a hybrid code-ized method which combined a motion compensation and conversion coding It is the transformer code method changed and outputted at other image streams from which syntax differs. A syntax parameter required for transformer codes including input image stream lost-motion vector information is extracted. An extracted parameter is changed and saved so that syntax of an output stream may be suited. Without performing inverse transformation to a transform coefficient if needed, thin out a bidirectional prediction frame from an input image stream, and time amount resolution is reduced. Reduce image size of this image stream that time amount resolution reduced, without performing inverse transformation to a transform coefficient, and space resolution is reduced. It is characterized by generating and outputting an image stream from which the time of an input, space resolution, or time amount resolution differs under right syntax from this acquired transform coefficient information and this saved syntax parameter information.

[0013] Moreover, it is characterized by amending image deterioration of this image stream that time amount and space resolution reduced, without performing inverse transformation to a transform coefficient.

[0014] (Invention of a transformer coder) An input image stream by which compression coding was carried out with a hybrid code-ized method which combined a motion compensation and conversion coding A means to be the transformer coder changed and outputted at other image streams from which syntax differs, and to extract a syntax parameter required for transformer codes including input image stream lost-motion vector information, A means to change and save an extracted parameter so that syntax of an output stream may be suited, A means to thin out a bidirectional prediction frame from an input image stream, and to reduce time amount resolution, without performing inverse transformation to a transform coefficient if needed, A means to reduce image size of this image stream that time amount resolution reduced, and to reduce space resolution, without performing inverse transformation to a transform coefficient, It is characterized by having a means which generates and outputs an image stream from which the time of an input, space resolution, or time amount resolution differs under right syntax from this acquired transform coefficient information and this saved syntax parameter information.

[0015] Moreover, it is characterized by having a means to amend image deterioration of this image stream that time amount and space resolution reduced, without performing inverse transformation to a transform coefficient.

[0016] (Invention of a record medium) An input image stream by which compression coding was carried out with a hybrid code-ized method which combined a motion compensation and conversion coding It is the record medium which recorded a program of a transformer code method changed and outputted on other image streams from which syntax differs. A process in which a syntax parameter required for

transformer codes including input image stream lost-motion vector information is extracted, A process in which an extracted parameter is changed and saved so that syntax of an output stream may be suited, A process in which thin out a bidirectional prediction frame from an input image stream, and time amount resolution is reduced, without performing inverse transformation to a transform coefficient if needed, A process in which reduce image size of this image stream that time amount resolution reduced, and space resolution is reduced, without performing inverse transformation to a transform coefficient, A process which generates and outputs an image stream from which the time of an input, space resolution, or time amount resolution differs under right syntax from this acquired transform coefficient information and this saved syntax parameter information, It is characterized by recording a program for performing a computer on a record medium which this computer can read.

[0017] Moreover, it is characterized by having a process which amends image deterioration of this image stream that time amount and space resolution reduced, without performing inverse transformation to a transform coefficient.

[0018]

[Embodiment of the Invention] Drawing 1 is a block block diagram of a transformer coder which shows the operation gestalt of this invention and does not have a drift error correction function. First, the input to this equipment is the input image stream 1 by which compression coding was carried out by the hybrid code-ized method which combined a motion compensation and conversion coding. There are MPEG-2 and MPEG-1 as an example of a concrete stream, and a discrete cosine transform (DCT) is used for conversion coding in these cases.

[0019] All the original methods that combined a motion compensation and other conversion coding methods (discrete sign conversion (DST), a Hadamard transform, haar conversion, slant conversion, KL conversion) besides these standard methods are contained in the input image stream 1.

[0020] Moreover, the output from this equipment is the output image stream 8 by which image size and frame rates were reduced rather than the input image stream. As an example of a concrete stream, MPEG-4, MPEG-2, MPEG-1, H.261, and H.263 grade are contained. The original format based on the same various conversion as an input is also included in the output image stream 8.

[0021] The flow of transformer code processing of this operation gestalt is explained. The variable-length decode of the input image stream 1 is first carried out in the variable-length decryption section 2. Furthermore, the syntax data 9 which is needed for generation of an output stream is extracted from an input stream. Moreover, the picture type of each frame is distinguished here, when frame rates need to be reduced, a bidirectional prediction frame (B-picture) can be dropped, the image stream obtained as a result is inputted into the input-side reverse quantization section 3, and reverse quantization of the transform coefficient of each block is carried out.

[0022] Next, the transform coefficient of each block by which reverse quantization was carried out is inputted into the functional-block group A which reduces image size with translation data. The functional-block group A separates in the pretreatment transducer 4 and the space down sample section 5. The pretreatment transducer 4 changes a little with classes of input image stream. If the case where the field image of MPEG-2 is changed into frame images including H.263, H.261, and MPEG-4 is considered as an example, in the pretreatment transducer 4, the coefficient changed in frame DCT mode will be first changed into the coefficient of Field DCT. In order to ask for the transform coefficient of the block which is equivalent to the bottom field from block B1j which adjoins up and down within each macro block, and B-2j ( $1 \leq j \leq 2$ ), the following formulas (1) are applied.

[0023]

[Equation 1]

$$B_j^{\text{bottom}} = \text{DCT}(f_1) \cdot B_{1j} + \text{DCT}(f_2) \cdot B_{2j} (j=1, 2) \quad \dots (1)$$

$$\text{where } f_1 = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \quad f_2 = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

[0024] Here, the function DCT (f) which acts on blocks f1 and f2 expresses conversion by two-dimensional DCT to 8x8 blocks, and is specifically defined by the following formulas.

[0025]

[Equation 2]

$$\text{DCT}(f) = (1/\sqrt{2}) A \cdot f \cdot A^T \quad \dots (2)$$

$$\text{where } A_{m,n} = (1/\sqrt{2}) \cdot k_n \cdot \cos[m(n+1/2) \cdot \pi/8] (m, n=0, 1, 2, \dots, 7)$$

$$k_n = \begin{cases} 1/\sqrt{2} (n=0) \\ 1 (n=1, 2, \dots, 7) \end{cases}$$

[0026] Next, these DCT coefficients are inputted into the space down sample section 5, and the down sample of a perpendicular / level both directions is performed with a DCT coefficient. For example, it is performed as follows for realizing one half of down samples. A vertical down sample is taking out only field of one of the two, and is realized easily. On the other hand, a horizontal down sample unifies two blocks which adjoined each other horizontally by the following formula with a DCT coefficient.

[0027]

[Equation 3]

$$\left. \begin{aligned} B_j^{\text{bottom}} [0..7][4..7] &\leftarrow 0 (j=1, 2) \\ \bar{B}^{\text{bottom}} &= B_1^{\text{bottom}} \cdot \text{DCT}(f_1^T) + B_2^{\text{bottom}} \cdot \text{DCT}(f_2^T) \end{aligned} \right\} \quad \dots (3)$$

[0028] Here, the 1st formula of a formula (3) expresses the actuation which transposes the component of 4 - 7 train (that is, considerable in the right half of a block) of Block B<sub>j</sub>bottom to 0. Still more generally it is also possible by unifying a formula (3) for the horizontal block of N individual of arbitration by the same matrix operation to realize the down sample of 1-/N. Moreover, the same processing is possible also in a perpendicular direction.

[0029] The transform coefficient which returned to drawing 1 and was outputted from functional-block A is inputted into the output side quantization section 6, and is quantized again. Next, variable length coding of the transform coefficient quantized in the output side quantization section 6 is inputted and carried out to the variable-length-coding section 7.

[0030] On the other hand, the input-side syntax data 9 extracted in the variable-length decryption section 2 is changed in accordance with an output stream by the syntax transducer 10, and is outputted as output side syntax data 11. It is compounded with the transform coefficient to which this output side syntax data 11 was outputted from the output side quantization section 6 in the variable-length-coding section 7, and, finally the output image stream 8 is generated and outputted.

[0031] Drawing 2 is a block block diagram of a transformer coder which shows other operation gestalten of this invention and has a drift error correction function. First, the input to this equipment is the input image stream 1 by which compression coding was carried out like drawing 1 by the hybrid code-ized method which combined a motion compensation and conversion coding. The variable-length decode of the input image stream 1 is first carried out in the variable-length decryption section 2. Furthermore, the input stream lost-motion vector information 13 and the input-side syntax data 9 which is needed for generation of an output stream in addition to this are extracted. Moreover, the picture type of each frame is distinguished, and when frame rates need to be reduced, a bidirectional prediction frame (B1 picture)

can be dropped here. The image stream (it consists of I picture and P1 picture) dropped in B1 picture is inputted into the input-side reverse quantization section 3, and reverse quantization of the transform coefficient of each block is carried out.

[0032] Next, the transform coefficient of each block by which reverse quantization was carried out is inputted into the functional-block group A which reduces image size with translation data. The functional-block group A separates in the pretreatment transducer 4 and the space down sample section 5. Although the pretreatment transducer 4 changes a little with classes of input image stream, it is the same as that of the case of drawing 1 about details.

[0033] Next, these DCT coefficients are inputted into the space down sample section 5, and the down sample of a perpendicular / level both directions is performed with a DCT coefficient. It is the same as that of the case of drawing 1 too for details.

[0034] Now, after the amendment signal 19 from the feedback loop is added, the transform coefficient 18 after scale conversion is inputted into the output side quantization section 6, and is quantized again. The output signal 20 of the output side quantization section 6 branches, while variable length coding of one side is inputted and carried out to the variable-length-coding section 7, the output side syntax data 11 by which syntax conversion was carried out by the syntax transducer 10 in accordance with the output stream is compounded, and the output image stream 8 is generated and outputted here.

[0035] Another side of an output signal 20 is inputted into the feedback loop B. First, it is inputted into the output side reverse quantization section 17, and reverse quantization is carried out. The coefficient 21 outputted from the output side reverse quantization section 17 takes difference with the transform coefficient 18 after scale conversion.

[0036] Here, the acquired differential signal is an error (drift error) signal produced by re-quantization in the present frame. Then, this differential signal is added to the prediction coefficient 22 by which the motion compensation was carried out. A prediction coefficient 22 is a coefficient obtained by branching in the amendment signal 19 generated and outputted in the motion compensation section 25 in a frequency domain, since it is equivalent to the error signal accumulated by the before frame, the total drift error accumulated by the present frame is acquired, and this error signal is accumulated in frame memories 23 and 24 here.

[0037] It moves and the motion vector which is needed in the motion compensation section 25 in a frequency domain uses the thing which was beforehand extracted from the input stream and which carried out scale conversion in accordance with image size by the scale transducer 15 from vector data 13. Moreover, in order to make a prediction image exact, in the motion compensation section 25 in a frequency domain, the image reduced to the frame memory 23 from the bottom field in the image reduced from the top field is saved at a frame memory 24, and both sides are used as a reference image.

[0038] In addition, the program for making a computer etc. perform the procedure thru/or computational algorithm of processing is recorded on a memory card, MO, CD, DVD, etc., and this invention can distribute it. [ example / of an operation gestalt / of this invention shown in drawing 1 or drawing 2 ]

[ the record medium which this computer can read, for example, a floppy (registered trademark) disk, ]

[0039]

[Effect of the Invention] It becomes possible to change into different-species image data, without transforming inversely also at once the transform coefficient of the frequency domain by which conversion coding was carried out according to this invention the above passage. Thereby, it is effective in the ability to shorten the transformer code processing time.

[0040] Moreover, there is an effect which can improve the image quality after a transformer code, without transforming inversely also at once the transform coefficient of the frequency domain by which conversion coding was carried out by adding the feedback loop for drift error correction.

---

[Translation done.]





## 【特許請求の範囲】

【請求項1】 動き補償と変換符号化を組み合わせたハイブリッド符号化方式により圧縮符号化された入力映像ストリームを、シンタックスの異なる他の映像ストリームに変換・出力するトランスコード方法であって、入力映像ストリームから動きベクトル情報をはじめとするトランスコードに必要なシンタックスパラメータを抽出し、抽出したパラメータを出力ストリームのシンタックスに適合するように変換・保存し、必要に応じて変換係数に逆変換を施すことなく入力映像ストリームから双方向予測フレームを間引いて時間解像度を低減し、変換係数に逆変換を施すことなく時間解像度の低減した該映像ストリームの画像サイズを縮小して空間解像度を低減し、得られた該変換係数情報と保存された該シンタックスパラメータ情報から、入力時と空間解像度または時間解像度が異なる映像ストリームを正しいシンタックスのもとで生成・出力する、ことを特徴とする映像ストリーム間のトランスコード方法。

【請求項2】 請求項1記載のトランスコード方法において、変換係数に逆変換を施すことなく時間および空間解像度の低減した該映像ストリームの画像劣化を補正することを特徴とする映像ストリーム間のトランスコード方法。

【請求項3】 動き補償と変換符号化を組み合わせたハイブリッド符号化方式により圧縮符号化された入力映像ストリームを、シンタックスの異なる他の映像ストリームに変換・出力するトランスコード方法であって、入力映像ストリームから動きベクトル情報をはじめとするトランスコードに必要なシンタックスパラメータを抽出する手段と、抽出したパラメータを出力ストリームのシンタックスに適合するように変換・保存する手段と、必要に応じて変換係数に逆変換を施すことなく入力映像ストリームから双方向予測フレームを間引いて時間解像度を低減する手段と、変換係数に逆変換を施すことなく時間解像度の低減した該映像ストリームの画像サイズを縮小して空間解像度を低減する手段と、得られた該変換係数情報と保存された該シンタックスパラメータ情報から、入力時と空間解像度または時間解像度が異なる映像ストリームを正しいシンタックスのもとで生成・出力する手段と、を備えたことを特徴とするトランスコード方法。

【請求項4】 請求項1記載のトランスコード方法において、変換係数に逆変換を施すことなく時間および空間解像度の低減した該映像ストリームの画像劣化を補正する手段を

備えたことを特徴とするトランスコード方法。

【請求項5】 動き補償と変換符号化を組み合わせたハイブリッド符号化方式により圧縮符号化された入力映像ストリームを、シンタックスの異なる他の映像ストリームに変換・出力するトランスコード方法のプログラムを記録した記録媒体であって、入力映像ストリームから動きベクトル情報をはじめとするトランスコードに必要なシンタックスパラメータを抽出する過程と、抽出したパラメータを出力ストリームのシンタックスに適合するように変換・保存する過程と、必要に応じて変換係数に逆変換を施すことなく入力映像ストリームから双方向予測フレームを間引いて時間解像度を低減する過程と、変換係数に逆変換を施すことなく時間解像度の低減した該映像ストリームの画像サイズを縮小して空間解像度を低減する過程と、得られた該変換係数情報と保存された該シンタックスパラメータ情報から、入力時と空間解像度または時間解像度が異なる映像ストリームを正しいシンタックスのもとで生成・出力する過程と、をコンピュータに実行させるためのプログラムを、該コンピュータが読み取り可能な記録媒体に記録したことを特徴とするトランスコード方法のプログラムを記録した記録媒体。

【請求項6】 請求項5記載の記録媒体において、変換係数に逆変換を施すことなく時間および空間解像度の低減した該映像ストリームの画像劣化を補正する過程を備えたことを特徴とするトランスコード方法のプログラムを記録した記録媒体。

## 【発明の詳細な説明】

## 【0001】

【発明の属する技術分野】本発明は映像符号化技術に属し、すでにあるビットレートで符号化された映像データを、時間又は空間解像度の異なる、より低いビットレートの映像データに再符号化（トランスコード）する方法およびトランスコードに関するものである。

## 【0002】

【従来の技術】動き補償と変換符号化を組み合わせたハイブリッド符号化方式で圧縮符号化された映像データを、時間及び空間解像度の異なる低ビットレートの映像データに再符号化する方法には、

(1) 文献「Susie J. Wee, John G. Apostolopoulos, Nick Feamster, "Field-to-Frame Transcoding with Spatial and Temporal Downsampling", ICIP'99.1999」等の前例がある。

【0003】一方、変換係数に逆変換を施すことなく、映像データの再符号化を行う方法には、

(2) 文献「西村 敏、笠井裕之、高屋和幸、亀山 渉、榎 直昭、花村 剛、富永英義、「リアルタイム MPEG-2 ビデオトランスコードソフトウェアの開発」、情処研報, vol. AVM25-5, pp. 25-30. 1999」、

(3) 文献「Pedro A. A. Assuncao, Mohammed Ghanbari, "

A Frequency-Domain Video Transcoder for Dynamic Bit-Rate Reduction of MPEG-2 Bit Streams", IEEE Trans. CSVT, vol. 8, pp. 953-967, Dec. 1998」等の前例がある。

【0004】

【発明が解決しようとする課題】上記の(1)の方式では、時間及び空間解像度の異なる映像データへの再符号化を行うため、圧縮された映像データを逆変換によりいったん伸長した後、画素領域で画像サイズを縮小するとともに、再度、動き補償と変換符号化による符号化を行う、という冗長な構成を取るため、トランスコード処理時間が増大するという問題があった。

【0005】また、このようなトランスコードに伴う処理時間の短縮を目的とした、逆変換を伴わない変換方式の提案(前記の(2)および(3))は、その適用領域はMPEG-2→MPEG-2等の同種ストリーム間で、時間及び空間解像度を変えないトランスコードに限られており、時間又は空間解像度の低減を伴う異種ストリームへの再符号化については提案されていなかった。

【0006】本発明の目的は、トランスコード処理時間を短縮して異種ストリームへの再符号化ができるトランスコード方法およびトランスコード並びにトランスコード方法のプログラムを記録した記録媒体を提供することにある。

【0007】

【課題を解決するための手段】本発明は、上記の課題を解決するため、以下に述べるような手段を採用している。

【0008】まず、ブロックごとに変換された変換係数を、逆変換を行わずに統合・縮小する機能ブロック群を用意する。本装置に入力した映像データは、可変長復号及び逆量子化を通して、変換係数ブロックとして該機能ブロック群に入力される。ここで、隣接する複数ブロックを統合した新たなブロックを生成することにより、ブロックサイズを縮小することなく画像全体の空間解像度低減を実現する。空間解像度が縮小された映像データは該機能ブロック群から出力され、目標ビットレートに応じて(通常は粗く)再量子化され、必要なシンタックスデータを追加・合成することにより、本装置から最終的に出力する映像ストリームを作成する。

【0009】このような処理により、変換係数の逆変換は一度も行わないことで、トータルの処理時間を短縮する。

【0010】また、一般に再量子化に関してはドリフト誤差に伴う画像劣化が避けられないため、本発明では、必要に応じて周波数データのままだりフト誤差を補正するフィードバックループを追加できるようにする。本フィードバックループには、周波数データのままだり補償を行う機能ブロック群が含まれ、トランスコード後の画質を向上させる。

【0011】以上のことから、本発明は、以下のトラン

スコード方法およびトランスコード並びにトランスコード方法を記録した記録媒体を特徴とする。

【0012】(トランスコード方法の発明)動き補償と変換符号化を組み合わせたハイブリッド符号化方式により圧縮符号化された入力映像ストリームを、シンタックスの異なる他の映像ストリームに変換・出力するトランスコード方法であって、入力映像ストリームから動きベクトル情報をはじめとするトランスコードに必要なシンタックスパラメータを抽出し、抽出したパラメータを出力ストリームのシンタックスに適合するように変換・保存し、必要に応じて変換係数に逆変換を施すことなく入力映像ストリームから双方向予測フレームを間引いて時間解像度を低減し、変換係数に逆変換を施すことなく時間解像度の低減した該映像ストリームの画像サイズを縮小して空間解像度を低減し、得られた該変換係数情報と保存された該シンタックスパラメータ情報から、入力時と空間解像度または時間解像度が異なる映像ストリームを正しいシンタックスのもとで生成・出力することを特徴とする。

【0013】また、変換係数に逆変換を施すことなく時間および空間解像度の低減した該映像ストリームの画像劣化を補正することを特徴とする。

【0014】(トランスコードの発明)動き補償と変換符号化を組み合わせたハイブリッド符号化方式により圧縮符号化された入力映像ストリームを、シンタックスの異なる他の映像ストリームに変換・出力するトランスコード方法であって、入力映像ストリームから動きベクトル情報をはじめとするトランスコードに必要なシンタックスパラメータを抽出する手段と、抽出したパラメータを出力ストリームのシンタックスに適合するように変換・保存する手段と、必要に応じて変換係数に逆変換を施すことなく入力映像ストリームから双方向予測フレームを間引いて時間解像度を低減する手段と、変換係数に逆変換を施すことなく時間解像度の低減した該映像ストリームの画像サイズを縮小して空間解像度を低減する手段と、得られた該変換係数情報と保存された該シンタックスパラメータ情報から、入力時と空間解像度または時間解像度が異なる映像ストリームを正しいシンタックスのもとで生成・出力する手段とを備えたことを特徴とする。

【0015】また、変換係数に逆変換を施すことなく時間および空間解像度の低減した該映像ストリームの画像劣化を補正する手段を備えたことを特徴とする。

【0016】(記録媒体の発明)動き補償と変換符号化を組み合わせたハイブリッド符号化方式により圧縮符号化された入力映像ストリームを、シンタックスの異なる他の映像ストリームに変換・出力するトランスコード方法のプログラムを記録した記録媒体であって、入力映像ストリームから動きベクトル情報をはじめとするトランスコードに必要なシンタックスパラメータを抽出する過程と、抽出したパラメータを出力ストリームのシンタッ

クスに適合するように変換・保存する過程と、必要に応じて変換係数に逆変換を施すことなく入力映像ストリームから双方向予測フレームを間引いて時間解像度を低減する過程と、変換係数に逆変換を施すことなく時間解像度の低減した該映像ストリームの画像サイズを縮小して空間解像度を低減する過程と、得られた該変換係数情報と保存された該シンタックスパラメータ情報から、入力時と空間解像度または時間解像度が異なる映像ストリームを正しいシンタックスのもとで生成・出力する過程と、をコンピュータに実行させるためのプログラムを、該コンピュータが読み取り可能な記録媒体に記録したことを特徴とする。

【0017】また、変換係数に逆変換を施すことなく時間および空間解像度の低減した該映像ストリームの画像劣化を補正する過程を備えたことを特徴とする。

【0018】

【発明の実施の形態】図1は、本発明の実施形態を示し、ドリフト誤差訂正機能を持たないトランスコードのブロック構成図である。まず、本装置への入力は、動き補償と変換符号化を組み合わせたハイブリッド符号化方式で圧縮符号化された入力映像ストリーム1である。具体的なストリームの例としてMPEG-2、MPEG-1があり、これらの場合、変換符号化には離散コサイン変換(DCT)が使用される。

【0019】入力映像ストリーム1には、これらの標準方式以外にも、動き補償と他の変換符号化方式(離散サイン変換(DST)、アダマル変換、ハール変換、スラント変換、KL変換)を組み合わせたあらゆる独自方式が含まれる。

【0020】また、本装置からの出力は、入力映像ストリームよりも画像サイズ及びフレームレートが削減され

た出力映像ストリーム8である。具体的なストリームの例として、MPEG-4、MPEG-2、MPEG-1、H.261、H.263等が含まれる。出力映像ストリーム8には、入力と同様の各種変換に基づく独自フォーマットも含まれる。

【0021】本実施形態のトランスコード処理の流れを説明する。入力映像ストリーム1は、まず可変長復号化部2で可変長復号される。さらに、入力ストリームから出力ストリームの生成に必要なシンタックスデータ9が抽出される。また、ここでは各フレームのピクチャタイプが判別され、フレームレートを削減する必要がある場合には双方向予測フレーム(B-ピクチャ)をドロップすることができ、その結果得られる映像ストリームは入力側逆量子化部3に入力され、各ブロックの変換係数が逆量子化される。

【0022】次に、逆量子化された各ブロックの変換係数は、変換データのまま画像サイズの縮小を行う機能ブロック群Aに入力される。機能ブロック群Aは、前処理変換部4と空間ダウンサンプル部5に別れる。前処理変換部4は、入力映像ストリームの種類により若干異なる。一例として、MPEG-2のフィールド画像をH.263、H.261、MPEG-4をはじめとするフレーム画像に変換する場合を考えると、まず前処理変換部4において、フレームDCTモードで変換されている係数がフィールドDCT相当の係数に変換される。各マクロブロック内で上下に隣接するブロック $B_{1j}$ 、 $B_{2j}$ ( $j=1, 2$ )からボトムフィールドに相当するブロックの変換係数を求めるには、例えば以下の式(1)が適用される。

【0023】

【数1】

$$B_j^{\text{bottom}} = \text{DCT}(f_1) \cdot B_{1j} + \text{DCT}(f_2) \cdot B_{2j} (j=1, 2) \quad \dots (1)$$

$$\text{where } f_1 = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \quad f_2 = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

【0024】ここで、ブロック $f_1$ 、 $f_2$ に作用する関数DCT( $f$ )は、 $8 \times 8$ ブロックに対する2次元DCTによる変換を表し、具体的には以下の式で定義される。

$$\text{DCT}(f) = (1/2) A \cdot f \cdot A^T \quad \dots (2)$$

$$\text{where } A_{m,n} = (1/2) \cdot k_n \cdot \cos[m(n+1/2) \cdot \pi/8] (m, n=0, 1, 2, \dots, 7)$$

$$k_n = \begin{cases} 1/\sqrt{2} (n=0) \\ 1 (n=1, 2, \dots, 7) \end{cases}$$

【0026】次に、これらのDCT係数は空間ダウンサンプル部5に入力され、DCT係数のまま垂直/水平両

方向のダウンサンプルが行われる。例えば、 $1/2$ のダウンサンプルを実現するには次のようにする。垂直方向

のダウンサンプルは片方のフィールドのみ取り出すことで、容易に実現される。一方、水平方向のダウンサンプルは、水平方向に隣り合った2つのブロックを、例えば

$$\left. \begin{aligned} B_j^{\text{bottom}}[0..7][4..7] &\leftarrow 0 (j=1, 2) \\ \tilde{B}^{\text{bottom}} &= B_1^{\text{bottom}} \cdot \text{DCT}(f_1^T) + B_2^{\text{bottom}} \cdot \text{DCT}(f_2^T) \end{aligned} \right\} \dots (3)$$

【0028】ここで、式(3)の第1式は、ブロック $B_j^{\text{bottom}}$ の4～7列(すなわちブロックの右半分に相当)の成分を0に置き換える操作を表す。さらに一般には、任意のN個の水平方向ブロックを式(3)を同様の行列演算で統合することにより、 $1/N$ のダウンサンプルを実現することも可能である。また、同様の処理は垂直方向においても可能である。

【0029】図1に戻って、機能ブロックAから出力された変換係数は、出力側量子化部6に入力され、再び量子化される。出力側量子化部6で量子化された変換係数は、次に可変長符号化部7に入力されて可変長符号化される。

【0030】一方、可変長復号化部2で抽出された入力側シンタックスデータ9は、シンタックス変換部10で出力ストリームにあわせて変換され、出力側シンタックスデータ11として出力される。該出力側シンタックスデータ11は、可変長符号化部7で出力側量子化部6から出力された変換係数と合成され、最終的に出力映像ストリーム8が生成・出力される。

【0031】図2は、本発明の他の実施形態を示し、ドリフト誤差訂正機能を持つトランスコードのブロック構成図である。まず、本装置への入力は、図1と同様に、動き補償と変換符号化を組み合わせたハイブリッド符号化方式で圧縮符号化された入力映像ストリーム1である。入力映像ストリーム1は、まず可変長復号化部2で可変長復号される。さらに、入力ストリームから動きベクトル情報13とその他出力ストリームの生成に必要な入力側シンタックスデータ9が抽出される。また、ここでは、各フレームのピクチャタイプが判別され、フレームレートを削減する必要がある場合には双方向予測フレーム(B-ピクチャ)をドロップすることができる。B-ピクチャをドロップされた映像ストリーム(I-ピクチャとP-ピクチャから成る)は、入力側逆量子化部3に入力され、各ブロックの変換係数が逆量子化される。

【0032】次に、逆量子化された各ブロックの変換係数は、変換データのまま画像サイズの縮小を行う機能ブロック群Aに入力される。機能ブロック群Aは、前処理変換部4と空間ダウンサンプル部5に別れる。前処理変換部4は、入力映像ストリームの種類により若干異なるが、詳細については図1の場合と同様である。

【0033】次に、これらのDCT係数は空間ダウンサンプル部5に入力され、DCT係数のまま垂直/水平両方向のダウンサンプルが行われる。詳細はやはり図1の

次の式によりDCT係数のまま統合する。

【0027】

【数3】

場合と同様である。

【0034】さて、スケール変換後の変換係数18は、フィードバックループからの補正信号19を加算された後、出力側量子化部6に入力され、再び量子化される。出力側量子化部6の出力信号20は分岐され、一方は可変長符号化部7に入力されて可変長符号化されるとともに、ここで、出力ストリームにあわせてシンタックス変換部10でシンタックス変換された出力側シンタックスデータ11が合成され、出力映像ストリーム8が生成・出力される。

【0035】出力信号20のもう一方はフィードバックループBに入力される。まず、出力側逆量子化部17に入力され、逆量子化される。出力側逆量子化部17から出力された係数21はスケール変換後の変換係数18との差分を取る。

【0036】ここで、得られた差分信号は、現フレームにおいて再量子化により生じた誤差(ドリフト誤差)信号である。引き続き、本差分信号は動き補償された予測係数22に加算される。予測係数22は、周波数領域内動き補償部25で生成・出力された補正信号19を分岐して得られる係数であり、前フレームまでに蓄積された誤差信号に相当することから、ここでは現フレームまでに蓄積されたトータルのドリフト誤差が得られ、本誤差信号がフレームメモリ23、24に蓄積される。

【0037】周波数領域内動き補償部25において必要となる動きベクトルは、あらかじめ入力ストリームから抽出しておいた動きベクトルデータ13から、スケール変換部15で画像サイズにあわせてスケール変換したものを使用する。また、予測画像を正確なものとするため、周波数領域内動き補償部25では、トップフィールドから縮小された画像をフレームメモリ23に、ボトムフィールドから縮小された画像をフレームメモリ24に保存して、双方を参照画像として利用する。

【0038】なお、本発明は、図1または図2に示した本発明の実施形態例での処理の手順ないし計算アルゴリズムをコンピュータ等に行わせるためのプログラムを該コンピュータが読み取り可能な記録媒体、例えばフロッピー(登録商標)ディスクやメモ리카ード、MO、CD、DVDなどに記録して配布することが可能である。

【0039】

【発明の効果】以上のとおり、本発明によれば、変換符号化された周波数領域の変換係数を一度も逆変換することなく、異種映像データへ変換することが可能となる。これにより、トランスコード処理時間を短縮できる効果

がある。

【0040】また、ドリフト誤差補正用のフィードバックループを追加することにより、変換符号化された周波数領域の変換係数を一度も逆変換することなく、トランスコード後の画像品質を向上できる効果がある。

【図面の簡単な説明】

【図1】本発明の実施形態を示すドリフト誤差訂正機能を持たないトランスコードのブロック構成図。

【図2】本発明の他の実施形態を示すドリフト誤差訂正機能を持つトランスコードのブロック構成図。

【符号の説明】

1…入力映像ストリーム

2…可変長復号化部

3…入力側逆量子化部

4…前処理変換部

5…空間ダウンサンプル部

6…出力側量子化部

7…可変長符号化部

10…シンタックス変換部

15…動きベクトルスケール変換部

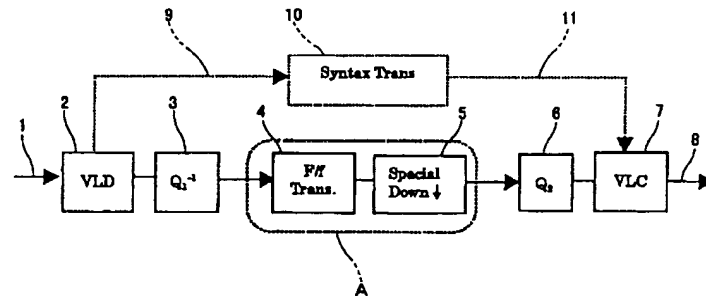
17…出力側量子化部

23…トップフィールド用フレームメモリ

24…ボトムフィールド用フレームメモリ

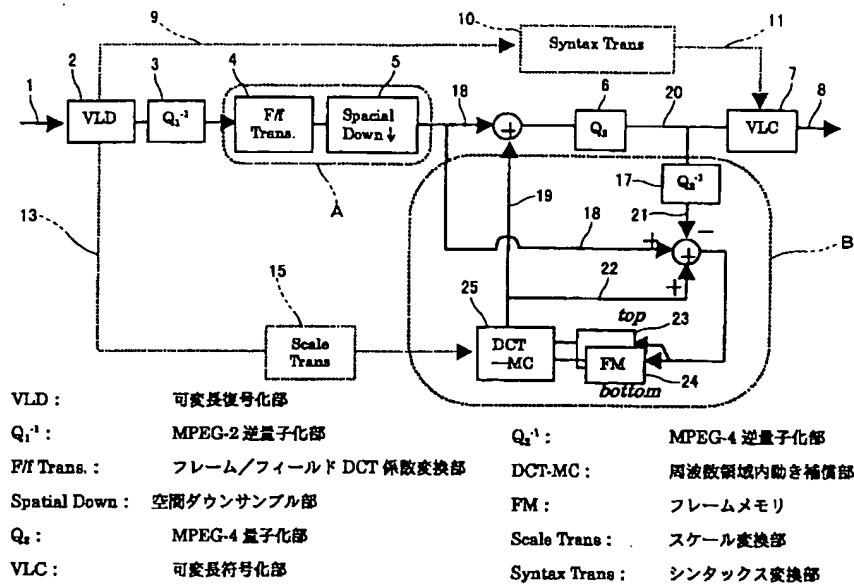
25…周波数領域内動き補償部

【図1】



VLD : 可変長復号化部  
 $Q_i^{-1}$  : MPEG-2 逆量子化部  
 F/H Trans : フレーム/フィールド DCT 係数変換部  
 Spatial Down : 空間ダウンサンプル部  
 $Q_2$  : MPEG-4 量子化部  
 VLD : 可変長符号化部  
 Syntax Trans : シンタックス変換部

【図 2】



フロントページの続き

(72)発明者 阪谷 徹

東京都千代田区大手町二丁目 3 番 1 号 日  
 本電信電話株式会社内

(72)発明者 小林 直樹

東京都千代田区大手町二丁目 3 番 1 号 日  
 本電信電話株式会社内

F ターム(参考) 5C059 KK01 KK33 KK41 MA05 MA23

MC11 MC38 ME01 ME15 NN01

NN28 RC28 RC31 TA45 TA53

TB08 TC37 TD13 UA02 UA05

5J064 AA03 BA09 BA16 BB13 BC08

BC16 BD01